Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper: PA—Precision Agriculture

Development of a stereovision sensing system for 3D crop row structure mapping and tractor guidance

Michio Kise[a], *Qin Zhang*[b,*]

[a]*Washington State University, Center for Precision Agricultural Systems, 24106 North Bunn Road, Prosser, WA 99350, USA*
[b]*Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, 1304 West Pennsylvania Avenue, Urbana, IL 61801, USA*

ARTICLE INFO

Article history:
Received 15 June 2007
Received in revised form
9 July 2008
Accepted 10 August 2008
Published online 24 September 2008

This research aimed to develop a field sensing system capable of performing three dimensional (3D) field mapping for measuring crop height and volume and detecting crop rows in 3D for reliable tractor guidance using one tractor-mounted stereocamera. The core of the dual-application field sensing system is a stereovision-based mapping method for creating 3D crop structure maps by means of estimating the motion of tractor-mount stereocamera and stitching constituent stereoimages progressively. By this approach, a feature point tracking algorithm was used to extract tractor motion indicating feature points from constituent stereoimages, and then feed the outcomes to a mathematical model of tractor dynamics to estimate tractor travelling speed and heading direction. Field validation test results indicated that the developed system could accurately create a 3D crop row structure map that can truthfully represent the spatial variability of crop height and volume from acquired stereoimages and can be used for reliably guiding the tractor by following crop rows.

1. Introduction

Crop growth condition monitoring and automated vehicle guidance are two major operations in mechanized precision agriculture. In recent years, commercial GPS-based guidance tractors and yield mapping systems have been made available by major agricultural machinery manufactures. Both applications are towards a similar goal—improving operation efficiency and productivity. However, such GPS-based technologies have a common limitation on obtaining local condition awareness which is often very important for performing efficient automated field operations.

Machine vision has been widely utilized as a condition awareness sensor because it offers the ability to instantly assess the target, and it does so in a non-destructive manner. Significant research has already been carried out in the machine vision-based guidance systems and field sensing systems (Reid and Searcy, 1987; Tillett *et al.*, 2001; Søgaard and Olsen, 2003). Field images acquired from vehicle-mounted cameras provided baseline data for both operations. However, previous studies have conducted these operations independently (Tillett and Hague, 1999; Tillett *et al.*, 2002; Benson *et al.*, 2003; Schleicher *et al.*, 2003). The research introduced herein aimed to explore a machine vision-based field recognition system capable of agricultural vehicle navigation and three dimensional (3D) crop mapping using one platform.

Significant research has also been conducted in remote sensing for agriculture (Yang *et al.*, 2001; Han *et al.*, 2002; Noh *et al.*, 2005). Many different spectral vegetation indices based on remote sensing data have been developed to estimate

* *Corresponding author.*
E-mail address: qinzhang@uiuc.edu (Q. Zhang).
1537-5110/$ – see front matter © 2008 IAgrE. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.biosystemseng.2008.08.001

Nomenclature

d_x	lateral camera position in vehicle coordinates, m
d_y	longitudinal camera position in vehicle coordinates, m
d_z	vertical camera position in vehicle coordinates, m
h	heading angle of the tractor, deg
l	translational distance of the tractor, m
m	motion vector of the tractor
N	number of pair of corresponding point
Q	error function
$\mathbf{R}_x(\theta)$	rotation matrix around X_V axis
$\mathbf{R}_Y(\theta)$	rotation matrix around Y_V axis
$\mathbf{R}_Z(\theta)$	rotation matrix around Z_V axis
X_C	X axis of camera coordinates
$X_CY_CZ_C$	camera coordinates
X_V	X axis of vehicle coordinates
$X_VY_VZ_V$	vehicle coordinates
Y_C	Y axis of camera coordinates
Y_V	Y axis of vehicle coordinates
Z_C	Z axis of camera coordinates
Z_V	Z axis of vehicle coordinates
x	x coordinate
y	y coordinate
z	z coordinates
α	tilt angle of the camera, deg
β	roll angle of the camera, deg
ϕ	rotational angle of the tractor, deg
γ	pan angle of the camera, deg
θ	rotational angle, deg

Subscripts

c	camera coordinates
i	image frame number
v	vehicle coordinates
0	image 0
1	image 1

biophysical parameters, such as the height, canopy area, volume, and leaf area index (LAI) (Thenkabail *et al.*, 2000; Bajwa and Tian, 2001; Goel *et al.*, 2003; Payero *et al.*, 2004). However, these methods could not determine 3D parameters of vegetation precisely without additional information, such as plant model, weather history, and soil type (Ji and Peters, 2007). In contrast, a stereovision system can obtain direct measurement of 3D vegetation structure along with the spectral information. The additional dimension of the scene provides some critical information for many agricultural applications, such as crop growth condition observation and physical parameter estimation (Lines *et al.*, 2001; He *et al.*, 2003; Rovira-Más *et al.*, 2005), as well as livestock 3D shape extraction (Wu *et al.*, 2004).

The objective of this research was to explore a dual application of a stereovision-based field condition awareness system that would be capable of performing 3D crop row structure map creation and detecting crop row locations for navigating a tractor in field. While the crop row detection method for tractor navigation has been developed previously (Kise *et al.*, 2005), and validated on separate field tests using the sole-function guidance system, this paper will focus on reporting the 3D crop row structure mapping method using the same hardware system developed for the tractor navigation application.

2. Materials and methods

2.1. *Hardware system*

A stereocamera (STH-MD1, VidereDesign, Menlo Park, USA) was installed in front of the test platform tractor (John Deere 7700, John Deere, Moline, USA) approximately 2.2 m above ground level and tilting downward. The tractor platform was modified by installing an electrohydraulic steering system in parallel to an existing hydraulic steering mechanism for implementing automated steering. The stereocamera used in this research consisted of two identical monochrome cameras with two image sensors aligned in parallel with each other at a 230 mm separation. Each monochrome camera had a 1.3 mega pixel CMOS imager and a 12 mm lens. With the given configuration, the stereocamera could capture 3D field images between approximately 2 m and 10 m in front of the moving tractor. An RTK-GPS and a fibre optical gyroscope (FOG) were installed on the tractor to record the travel path and direction of the vehicle for validation purposes.

2.2. *3D crop row structure mapping algorithm*

3D crop row structure map creation was performed by stitching consecutive images from a vehicle-mounted stereocamera progressively. Fig. 1 shows the basic procedures of 3D crop structure map creation. The first step was stereo image acquisition and processing that computes the disparity image to obtain 3D information of the field scene. It was followed by the 3D crop row image creation, in which a disparity image was converted to 3D elevation data of a small patch in front of the platform tractor with respect to the vehicle coordinates. The feature point extraction is an image process that extracts distinctive features such as edges by applying a neighbourhood operator to the image. The subsequent feature point tracking compares two adjacent images and finds corresponding feature points. The camera motion, that is the motion of the tractor platform, is then estimated from the result of the feature point tracking. Constituent images were then projected on the field image coordinates progressively, based on the vehicle motion obtained from the preceding process. All algorithms were developed in C++ environment (Microsoft VisualStudio.NET 2003, Redmond, USA).

2.3. *3D crop row image creation*

A 3D crop row image represents the elevation of a small patch in front of the platform vehicle, and is baseline data for both guidance and 3D crop row structure mapping applications. In the guidance application, the relative position of the vehicle to crop rows was measured by detecting crop rows based on the

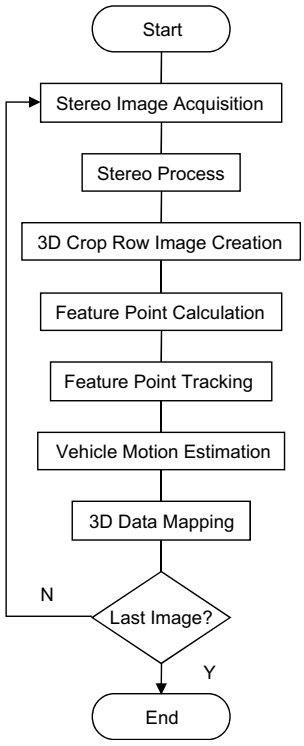

Fig. 1 – Flowchart of 3D crop row structure map creation.

height difference between crop rows and the ground in the 3D crop row image. In 3D crop structure mapping application, 3D crop row images collected from an entire field were integrated into one map. A 3D crop row image was created from a single disparity image. Fig. 2 shows an example of a 3D crop row image created from a stereo image taken in a soya bean field. As shown in Fig. 2(a), the original field scene image represented soya bean rows with crop height approximately 0.40 m and row spacing of 0.75 m. The disparity map of stereo image (Fig. 2(b)) encoded the depth of a scene, therefore, provided the baseline data for the 3D crop row image creation. The grey level of the 3D crop row image (Fig. 2(c)) indicates the height at the corresponding location in the field as a brighter pixel represents higher point. The resolution of the 3D crop row image was 150 pixels × 150 pixels, which was equivalent to a 3 m × 3 m square area in the field. With the particular stereo system arrangement, the height resolution of the map was approximately 5 mm.

A 3D crop row image creation algorithm was formulated by defining the vehicle coordinate system with its origin at the vehicle centre of gravity (CG), as illustrated in Fig. 3. The X_V and Y_V axes were defined as the vehicle lateral and longitudinal directions and the Z_V axis was defined as the vertical, respectively. To compensate for camera mount angle, a 3D point obtained from a disparity image with respect to the camera coordinates (x_c, y_c, z_c) was transformed to the vehicle coordinates (x_v, y_v, z_v) by

$$\begin{pmatrix} x_v \\ y_v \\ z_v \end{pmatrix} = \mathbf{R}_Z(\gamma)\mathbf{R}_X(-\alpha)\mathbf{R}_Y(\beta)\begin{pmatrix} x_c \\ z_c \\ -y_c \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \quad (1)$$

where α, β, and γ are the tilt, roll, and pan angles of the camera mount (d_x, d_y, d_z) is the camera position offset to the tractor CG. $\mathbf{R}_X$, $\mathbf{R}_Y$ and $\mathbf{R}_Z$ are rotation matrices around X_V, Y_V, and Z_V axes and are represented by

$$\mathbf{R}_X(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

$$\mathbf{R}_Y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$

$$\mathbf{R}_Z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where θ is a rotational angle.

The tilt, roll and pan angles were identified based on a stereo image: the platform tractor was set stationary on a flat surface and acquired a stereo image of the surface. A flat surface on which the vehicle was placed represents X_VY_V plane, therefore, the 3D data of the flat surface calibrated with the optimum tilt and roll angles have to minimize following equation:

$$E = \sum (z_v - \bar{z}_v)^2 \quad (2)$$

where $\bar{z}_v$ is the average of z_v. The value of z_v in Eq. (2) was calculated by Eq. (1) with $d_z = 0$. With $d_z = 0$, $\bar{z}_v$ is essentially the camera height d_z.

The tilt and roll angles that minimize E of Eq. (2) could be identified by some optimization algorithm. A gradient descent was used in this study (Jongen *et al.*, 2004). Gradient descent finds a local minimum of a function so it is important to have

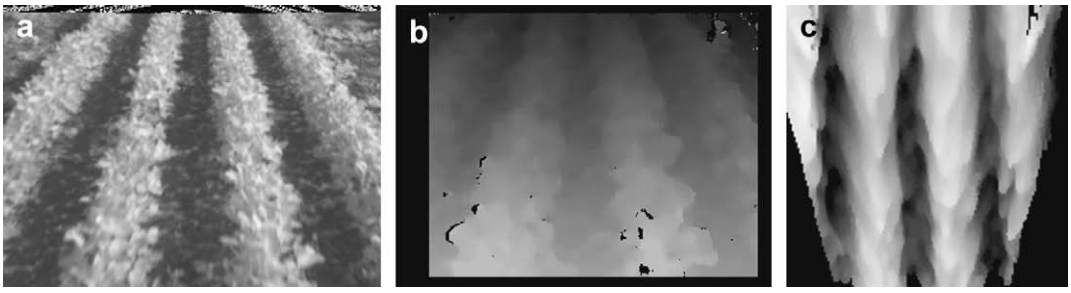

Fig. 2 – Creation of a 3D crop row image: (a) original image (left camera); (b) disparity image; and (c) resulting 3D crop row image. The grey level in (b) and (c) indicates the crop height: the brighter in colour, the higher in height.

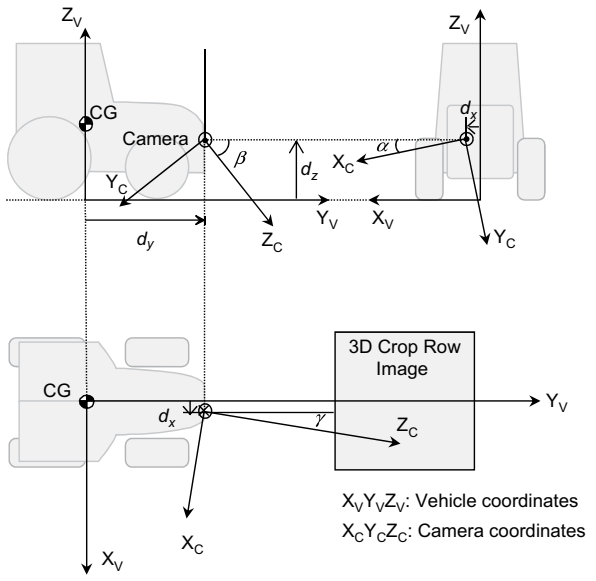

Fig. 3 – Camera installation and coordinate systems definition: CG; d_x, d_y and d_z, longitudinal, horizontal and vertical offsets of the camera location in vehicle coordinates; α, β and γ, roll, tilt and pan angles of camera installation.

a good initial value. However in this application, it would not be a problem because the camera mount angle could be measured manually with a decent accuracy. As a result, each parameter was identified as $\alpha = -0.5°$, $\beta = 41.5°$, $\gamma = 0.0°$, $d_x = 0.04$ m, $d_y = 2.65$ m, and $d_z = 2.22$ m.

2.4. *Feature point tracking*

The camera motion was estimated by means of tracking the feature points in two adjacent stereoimages. A feature point is defined as a pixel that has a distinctively different value from its neighbour pixels. Feature points can be detected by a feature point detector, which quantifies the distinctiveness of the pixel based on surrounding pixel values (Harris and Stephens, 1988; Shi and Tomashi, 1994). By searching for similar features among two adjacent images, corresponding points between two images can be determined. Feature point tracking algorithms are often found in machine vision applications such as face recognition (Moriyama *et al.*, 2006), motion detection (Poelman and Kanade, 1997) and image mosaicing (Kise and Zhang, 2008).

Feature points of this study were chosen based on the magnitude of the gradient. A gradient was calculated by applying a derivative filter to the pixel of interest and its neighbours. A 3×3 Sobel mask was used as the derivative filter (Gonzalez and Woods, 1992). Feature points found in one image were searched for in the following image. Each feature point found in preceding image (image 0) was compared with feature points in the following image (image 1) for their similarities to determine if they represent the same point. To compare their similarities, a feature point and its neighbour pixels in image 0 were warped to the image 1 coordinates. If image 0 was a 3D image and the geometric relationship between the two cameras (image 0 and 1) was known, it should be possible to reconstruct a hypothetical camera that had a view of image 1 from their geometric relationship (Scharstein, 1999).

If the two feature points represent the same point in the field, in theory the warped view from the image 0 is identical with the view of the image 1. The warped image is compared with an equal size window centred at the feature point in the image 1. The sum of the squared difference (SSD) of grey levels across the windows was used as an indicator of the similarity of the two windows.

$$S(x_0, y_0, d_x, d_y) = \sum_{i=x_0-w}^{x_0+w} \sum_{j=y_0-w}^{y_0+w} \{I_0(i,j) - I_1(i+d_x, j+d_y)\}^2 \quad (3)$$

where $S(x_0, y_0, d_x, d_y)$ is SSD between the window centred at (x_0, y_0) in image 0 and the window centred at $(x_0 + d_x, y_0 + d_y)$ in image 1, $2w + 1$ is the window size, $I_0(i, j)$, and $I_1(i, j)$ are the grey level of pixel (i, j) in image 0 and image 1, respectively. In this case, (x_0, y_0) in image 0 and $(x_0 + d_x, y_0 + d_y)$ in image 1 were the feature points to be compared for their matching. Obviously, the smallest S occurred at the best match of the points.

Fig. 4 shows an example of feature point tracking. A gradient was calculated at all pixels in the image, and 400 largest gradients were selected as the feature points of the image (Fig. 4(a)). As a result of the feature point tracking, 89 corresponding points were found in the consequent image (Fig. 4(b)).

2.5. *Tractor motion estimation*

It is necessary to define how to represent the camera motion to support the estimation process based on the corresponding feature points between two consequent images. Fig. 5 presents the definition of a motion vector in terms of the camera motion information extracted from two adjacent images based on a simple tractor dynamics model. This motion vector could represent both vehicle translational and rotational motions with reference to the vehicle coordinates at the image 0, as follows:

$$m = [l, \phi]^T \quad (4)$$

where m is a motion vector, l is the translational distance and ϕ is the rotational angle of the tractor.

The tractor dynamics model shown in Fig. 5 was designed under the assumption that the tractor travels on a flat surface so that the vertical translation and rotations about X_V and Y_V axes (pitching and rolling) of the tractor could be ignored. With corresponding points identified by the feature point tracking, the geometric relationship of a corresponding point between two images can be expressed by as follows:

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + l \begin{pmatrix} \sin\phi \\ \cos\phi \end{pmatrix} \quad (5)$$

where $[x_0, y_0]^T$ and $[x_1, y_1]^T$ are the location of the corresponding points in image 0 ($X_V^{(0)}Y_V^{(0)}$ coordinates) and image 1 ($X_V^{(1)}Y_V^{(1)}$ coordinates), respectively. Consider that ϕ and l are small, and ignore second-order terms; Eq. (5) becomes

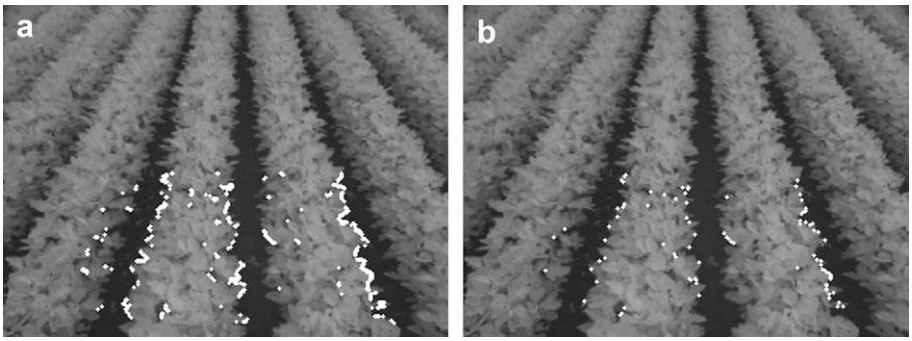

Fig. 4 – Feature point tracking: (a) 400 largest gradients of all pixels are selected as feature points and (b) as the result of the feature point tracking, 89 corresponding points are found in the adjacent image.

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_1 + \phi y_1 \\ -\phi x_1 + y_1 + l \end{pmatrix} \tag{6}$$

The motion vector can be calculated by means of a least squares method based on the multiple set of corresponding points.

$$Q = \sum_{k=0}^{N-1} \left\{ \left(x_{0,k} - x_{1,k} - \phi y_{1,k} \right)^2 + \left(y_{0,k} + \phi x_{1,k} - y_{1,k} - l \right)^2 \right\} \tag{7}$$

where Q is the error function to be minimized; $[x_{0,k}, y_{0,k}]^T$ and $[x_{1,k}, y_{1,k}]^T$ are pair of corresponding points found by the feature point tracking; and N is the number of pair of corresponding point found. $[l, \phi]^T$ that makes Q minimum is expressed by

$$\phi = \frac{N \sum_{k=0}^{N-1} \left\{ x_{1,k} y_{0,k} - y_{1,k} x_{0,k} \right\} + \sum_{k=0}^{N-1} x_{1,k} \sum_{k=0}^{N-1} (y_{1,k} - y_{0,k})}{\left(\sum_{k=0}^{N-1} x_{1,k} \right)^2 - N \sum_{k=0}^{N-1} \left\{ (x_{1,k})^2 + (y_{1,k})^2 \right\}} \tag{8}$$

$$l = \frac{\sum_{k=0}^{N-1} \left\{ (x_{1,k})^2 + (y_{1,k})^2 \right\} \sum_{k=0}^{N-1} (y_{1,k} - y_{0,k}) + \sum_{k=0}^{N-1} x_{1,k} \sum_{k=0}^{N-1} \left\{ x_{1,k} y_{0,k} - y_{1,k} x_{0,k} \right\}}{\left(\sum_{k=0}^{N-1} x_{1,k} \right)^2 - N \sum_{k=0}^{N-1} \left\{ (x_{1,k})^2 + (y_{1,k})^2 \right\}} \tag{9}$$

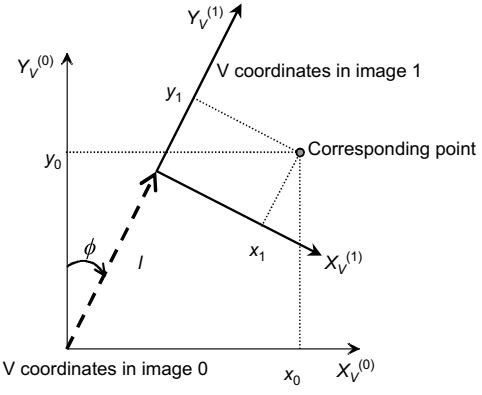

Fig. 5 – Tractor dynamics model: $X_V^{(0)}Y_V^{(0)}$, vehicle coordinates in image 0; $X_V^{(1)}Y_V^{(1)}$, vehicle coordinates in image 1; *f*, rotational angle of the tractor; *l*, translational distance of the tractor; and - - - - -, motion vector.

From Eqs. (5), (8) and (9), the travel path of the vehicle, namely the trajectory of the origin of vehicle coordinate system ($X_V Y_V Z_V$) with respect to field coordinates, could be represented in recursive form based on dead reckoning (Mizushima *et al.*, 2002):

$$\begin{pmatrix} ox_{t+1} \\ oy_{t+1} \end{pmatrix} = \begin{pmatrix} \cos\phi_t & -\sin\phi_t \\ \sin\phi_t & \cos\phi_t \end{pmatrix} \begin{pmatrix} ox_t \\ oy_t \end{pmatrix} + \begin{pmatrix} 0 \\ l_t \end{pmatrix}$$
$$oz_{t+1} = oz_t \tag{10}$$
$$h_{t+1} = h_t + \phi_t$$

with

$$ox_0 = 0$$
$$oy_0 = 0$$
$$oz_0 = 0$$
$$h_0 = 0$$

where *t* is image frame number (ox, oy, oz) is the origin of vehicle coordinate system with respect to map coordinates, and *h* is the angle between the map coordinate system and vehicle coordinates system. Each crop row structure image could be merged in a stepwise fashion to form an entire 3D field map. If there was an overlapping area between two images, the average of overlapping pixels was used.

3. Results and discussions

3.1. *Test field*

Series of stereoimages were acquired in a soya bean field for validating the developed algorithm. The test field was located at an experimental farm at University of Illinois at Urbana-Champaign. It was divided into three plots in terms of the day of planting. Each plot had 40 rows, 80 m long of soya beans with 0.75 m inter row spacing. On the day the images were collected, the crops in the three plots were at 44, 52, and 65 days after planting (DAP), and consequently, the approximate crop heights were 0.55 m, 0.70 m, and 0.80 m, respectively. The tractor travelled 10 paths in each plot to obtain 3D field

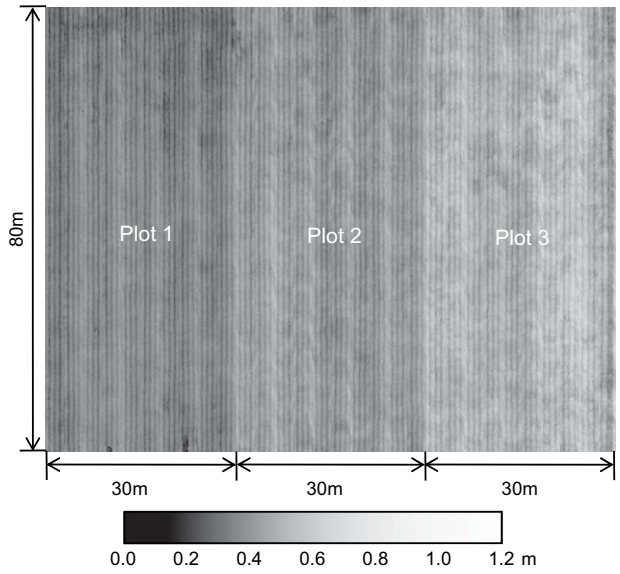

Fig. 6 – 3D crop row structure map created by stereovision-based mapping system; 90 m × 80 m entire 3D map consisting of three plots of 44, 52, and 65 DAPs. The grey level indicates the crop height: the brighter in colour, the higher in height.

images, resulting in a total of 30 paths for the entire field. The stereocamera recorded images at 5 frames s^{-1} with the tractor travelling at 1.0 m s^{-1}. The tractor was equipped with an RTK-GPS to record the paths for referencing. The height of plants was measured manually at 90 locations as ground truth data. Thirty single plants were selected randomly for each plot, and the height of the plant from the ground to the top of the main plant stem was measured by a ruler. The locations of the measured plants were also recorded by RTK-GPS.

3.2. *3D crop row structure map*

Fig. 6 shows a 3D soya bean crop row structure map created from the stereoimages. A grey level of the map indicates crop heights at the corresponding locations in the field with a brighter pixel representing a higher point. The spatial resolution of the map is 1.5 cm $pixel^{-1}$. The vertical direction (upward) corresponds to the northing.

The map of each single path was created individually, and then combined side-by-side to form the entire field map. The total 30 constituent paths comprise the entire map. Each path was wide enough to contain four rows. The images were always recorded from the north side of the field in order to align each single path map.

As expected, the height differences in three plots (44, 52, and 65 DAP) were apparent in the map: Plot 1 (44 DAP) was the lowest height, while plot 3 (65 DAP) was the highest in the field. The map also shows the intra-plot variability. Fig. 7(b) is the enlarged view of the portion of the map corresponding to a 2.7 m × 4.8 m rectangle region at the right edge of plot 3. As shown in the original stereo image (Fig. 7(a)), one row was distinctively smaller than the rest of the rows, as a result of being under-fertilized. The enlarged map demonstrated this size difference; the crops in the second row from the left in the map were darker (which means lower height) and thinner than those in the other three rows.

The enlarged map also shows that not only could the height of just one plant be measured but also its volume. For example, the total canopy volume of Fig. 7(b) could be calculated from simply summing all pixel values in the figure.

Fig. 8 shows the comparison of crop heights of 90 points estimated from the field 3D map against the ground truth data measured manually. The validation showed that the root mean squared (RMS) error between the 3D map values and the ground truth data was 0.04 m, with the maximum error of 0.09 m. This validation result proved that the 3D field mapping system developed in this research could provide centimetre-level crop plant height information with a high spatial resolution in the form of a panoramic field view.

3.3. *Tractor motion estimation*

The accuracy of travel speed and heading angle obtained by tractor motion estimation determines the mapping accuracy in terms of plant location (*XY* plane). Fig. 9 shows the

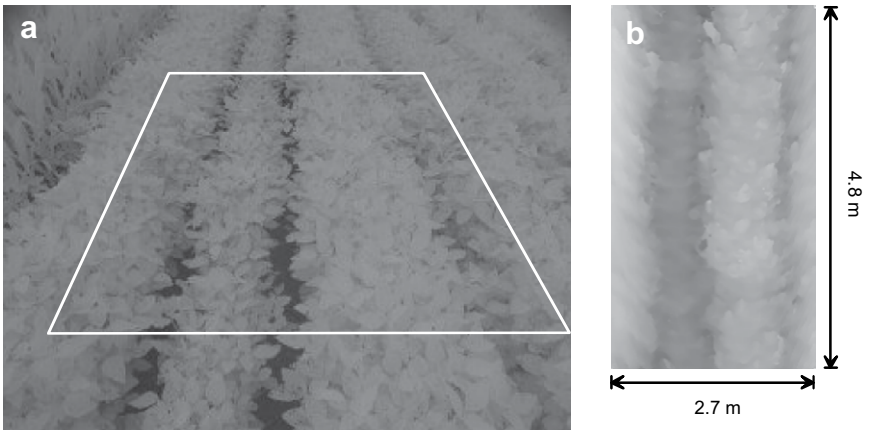

Fig. 7 – Enlarged view of the 3D crop row structure map: (a) the original field scene taken by the stereocamera and (b) enlarged portion (2.7 m × 4.8 m) of the 3D crop row structure map corresponding to the white rectangle area in Fig. 7(a). The grey level in (b) indicates the crop height: the brighter in colour, the higher in height.

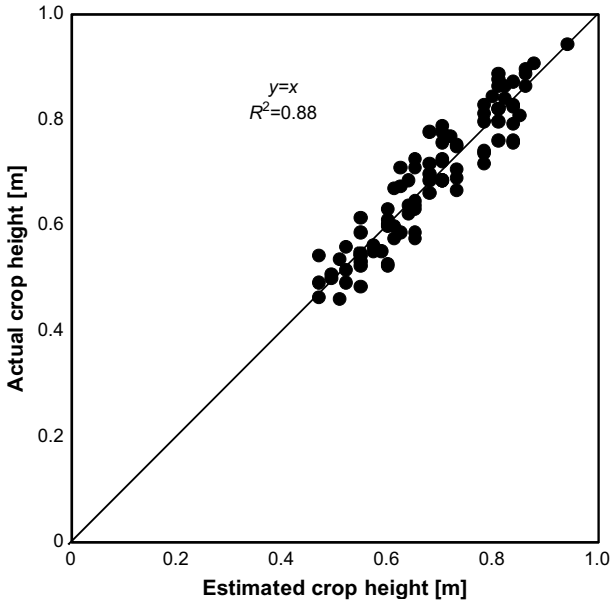

Fig. 8 – Comparison of the crop height from the 3D crop row structure map against the ground truth data.

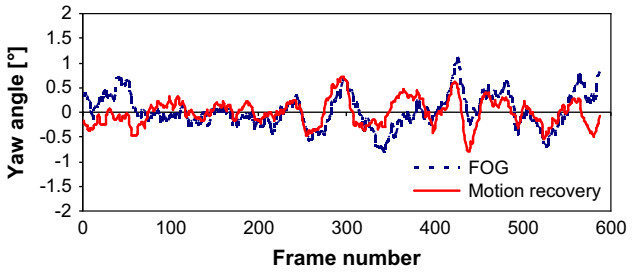

Fig. 10 – Performance of yaw angle estimation using the motion recovery algorithm.

estimated travelling speed of the tractor using Eq. (9) in comparison to the RTK-GPS data. The results indicated that the maximum error of speed estimation was within 0.30 m s^{-1} for tractor travelling speed range between 0.70 m s^{-1} and 1.10 m s^{-1}. Fig. 9 also revealed that the estimation errors were not uniformly distributed but tend to have a relatively constant offset for each trip. This offset, observed in almost all test runs, was presumably caused by the data acquisition program that didn't acquire the stereoimages at a constant rate—the image acquisition rate might periodically change during the data acquisition—resulting in the constant-offset-error of the travel speed estimation. The data indicated that 90% of the errors were distributed within a range of ±0.20 m s^{-1}, and the RMS error was 0.11 m s^{-1}.

Fig. 10 shows the comparison of the estimated yaw angle to the FOG sensor. In contrast to travel speed estimation, the estimated yaw angle did not result in any noticeable constant-offset-error. Such a difference could be attributed to the fact that the yaw angle estimation was a result of accumulation of angular changes between consecutive stereoimages, which made the resulting angle less sensitive to the irregular timing in image acquisition. 90% of the errors were distributed within a range of ±0.67° with the RMS error of 0.34°.

Very similar results were obtained from all three plots: the average RMS speed and yaw angle estimation errors were 0.11 m s^{-1} and 0.38° from plot 1, 0.13 m s^{-1} and 0.44° from plot 2, and 0.12 m s^{-1} and 0.37° from plot 3, respectively. The results proved that the motion recovery algorithm could reliably estimate both travel speed and yaw angle of a tractor at a sufficient accuracy for guidance applications.

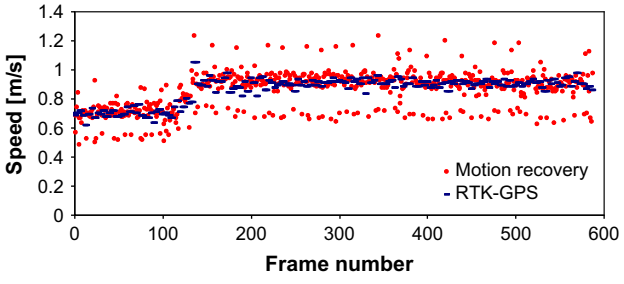

Fig. 9 – Performance of travel speed estimation using the motion recovery algorithm.

4. Conclusion

The development of 3D field mapping system for panoramic 3D crop row structure map creation was reported. Supported by the developed algorithm, a tractor-mounted stereocamera could acquire a stream of field stereoimages and create a 3D crop row structure map of an entire field using the acquired images. This map of crop row structure could not only be used to measure crop plant height and volume at a high spatial resolution (1.5 cm in horizontal plane and 5.0 mm in vertical axis) and a high accuracy (4.0 cm in plant height estimation), but also provide vehicle guidance information with sufficient accuracy (less than 0.13 m s^{-1} and 0.44° for travelling speed and yaw angle estimation). The results proved that this developed algorithm could be a very useful tool for many applications, e.g. plant growth condition monitoring by comparing two maps created from different dates and plant volume calculation for biomass productivity estimation. A study on addressing the use of GPS for compensating drift error could further improve the accuracy of system.

Acknowledgements

The material presented in this paper was based upon work supported partially by Illinois Council on Food and Agricultural Research (IDA CF 99 SI-36-1A) and USDA Hatch Funds (ILLU-10-352 AE). Japan Society for the Promotion of Science (JSPS) provides a Fellowship Fund to support Dr. Michio Kise conducting postdoctoral research at the University of Illinois at Urbana-Champaign. Any opinions, findings, and conclusions expressed in this publication are those of the authors and do not necessarily reflect the views of the University of Illinois, JSPS, Illinois CFAR and USDA.

REFERENCES

Bajwa S G; Tian L F (2001). Aerial CIR remote sensing for weed density mapping in a soybean field. Transactions of the ASAE, **44**(6), 1965–1974.

Benson E R; Reid J F; Zhang Q (2003). Machine vision-based guidance system for an agricultural small-grain harvester. Transactions of the ASAE, **46**(4), 1255–1264.

Goel P K; Landy J A; Patel R M; Viau A A; Miller J R (2003). Estimation of crop biophysical parameters through airborne and field hyperspectral remote sensing. Transactions of the ASAE, **46**(4), 1235–1246.

Gonzalez R C; Woods R E (1992). Digital Image Processing. Addison-Wesley Publishing Company, Reading, USA.

Han S; Hendrickson L; Ni B (2002). Comparison of satellite and aerial imagery for detecting leaf chlorophyll content in corn. Transactions of ASAE, **45**(4), 1229–1236.

Harris C; Stephens M (1988). A combined corner and edge detector. In: Proceedings of Fourth Alvey Vision Conference, pp 147–151, Manchester, UK.

He D X; Matsuura Y; Kozai T; Ting K C (2003). A binocular stereovision system for transplant growth variables analysis. Applied Engineering in Agriculture, **19**(5), 611–617.

Ji L; Peters A J (2007). Performance evaluation of spectral vegetation indices using a statistical sensitivity function. Remote Sensing of Environment, **106**(1), 59–65.

Jongen H T; Meer K; Triesch E (2004). Optimization Theory. Kluwer Academic Publishers, Boston, MA.

Kise M; Zhang Q (2008). Creating a panoramic field image using multi-spectral stereovision system. Computers and Electronics in Agriculture, **60**(1), 67–75.

Kise M; Zhang Q; Rovira–Más F (2005). A stereovision-based crop row detection method for tractor-automated guidance. Biosystems Engineering, **90**(4), 357–367.

Lines J A; Tillett R D; Ross L G; Chan D; Hockaday S; McFarlane N J B (2001). An automatic image-based system for estimating the mass of free-swimming fish. Computers and Electronics in Agriculture, **31**(2), 151–168.

Mizushima A; Noguchi N; Ishii K; Terao H (2002). Automatic navigation of the agricultural vehicle by the geomagnetic direction sensor and gyroscope. In: Proceedings of Automation Technology for Off-Road Equipment (ATOE), pp 204–211, Chicago, USA.

Moriyama T; Kanade T; Xiao J; Cohn J F (2006). Meticulously detailed eye region model and its application to analysis of facial images. IEEE Transactions on Pattern Analysis and Machine Intelligence, **28**(5), 738–752.

Noh H; Zhang Q; Han S; Shin B; Reum D (2005). Dynamic calibration and image segmentation methods for multispectral imaging crop nitrogen deficiency sensors. Transactions of the ASAE, **48**(1), 393–401.

Payero J O; Neale C M U; Wright J L (2004). Comparison of eleven vegetation indices for estimating plant height of alfalfa and grass. Applied Engineering in Agriculture, **20**(3), 385–393.

Poelman C J; Kanade T (1997). A paraperspective factorization method for shape and motion recovery. IEEE Transactions on Pattern Analysis and Machine Intelligence, **19**(3), 206–218.

Reid J F; Searcy S W (1987). Vision-based guidance of an agricultural tractor. IEEE Control Systems Magazine, **7**(12), 39–43.

Rovira-Más F; Zhang Q; Reid J F (2005). Creation of three-dimensional crop maps based on aerial stereoimages. Biosystems Engineering, **90**(3), 251–259.

Scharstein D (1999). View Synthesis Using Stereo Vision. In: Lecture Note in Computer Science. Springer-Verlag, Berlin, Germany.

Shi J; Tomashi C (1994). Good features to track. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR94) Seattle, USA.

Schleicher T D; Bausch W C; Delgado J A (2003). Low ground-cover filtering to improve reliability of the nitrogen reflectance index (NRI) for corn N status classification. Transactions of the ASAE, **46**(6), 1707–1711.

Søgaard H T; Olsen H J (2003). Determination of crop rows by image analysis without segmentation. Computers and Electronics in Agriculture, **38**(2), 141–158.

Thenkabail P S; Smith R B; De Pauw E (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. Remote Sensing of Environment, **71**(2), 158–182.

Tillett N D; Hague T (1999). Computer-vision-based hoe guidance for cereals – an initial trial. Journal of Agricultural Engineering Research, **74**(3), 225–236.

Tillett N D; Haugue T; Miles S J (2001). A field assessment of a potential method for weed and crop mapping on the basis of crop planting geometry. Computers and Electronics in Agriculture, **32**(3), 229–246.

Tillett N D; Hague T; Miles S J (2002). Inter-row vision guidance for mechanical weed control in sugar beet. Computers and Electronics in Agriculture, **33**(3), 163–177.

Wu J; Tillett R; McFarlane N; Ju X; Siebert J P; Schofield P (2004). Extracting the three-dimensional shape of live pigs using stereo photogrammetry. Computers and Electronics in Agriculture, **44**(3), 203–222.

Yang C; Bradford J M; Wiegand C L (2001). Airborne multispectral imagery for mapping variable growing conditions and yields of cotton, grain sorghum, and corn. Transactions of the ASAE, **44**(6), 1983–1994.